City of Albuquerque
Information Technology Services Division
Data Management

data.cabq.gov Core Metadata Requirements

Contact Information

Who is the contact for this dataset? The contact will be the City Employee who is accountable for the data provided in this dataset and con act as front-line support in the event of any questions about the data.

Name	Narong Joe Saraphon
Department/Division	Transit
Email	saraphon@cabq.com

What Does this Dataset Describe?

What is the name of this dataset? How should a user identify this dataset in any communication with contact above? Provide a shorter description of the Dataset that can act as a one-line summary of the dataset when dealing with stakeholders. Provide a longer description of the data that can be readily understood by non-technical users.

Dataset Title	abqbusesYYYYMM.txt
Short Description	Historical bus locations
Full Non-Technical Description	Bus locations originally gathered for use with near real-time reporting to smart phone app and wmb.unm.edu.

How Should this Dataset be Cited

How should external sources refer to this dataset in publications or documentation? Often this will simply be the URL and the data retrieved.

http://data.cabq.gov/transit/historical/abqbusesYYYYMM.txt

Does the Dataset Reflect a Particular Time Period?

Provide any date restrictions that may affect the validity of the data. The table fields are defined as follows:

Field	Definition
Start Date	Start date of the time period within which this data falls. Format: MM/DD/YYYY HH:MM:SS
End Date	End date of the time period within which this data falls. Format: MM/DD/YYYY HH:MM:SS
Dataset Refresh Interval	Time period between Dataset refreshes. Format: "nn [seconds\minutes\hours\days\weeks\months\years]" or the word "Static" if never refreshed.
Dataset Expiration Date	Date after which the data must be considered stale and no longer of sufficient utility (fit-for-purpose). Format: MM/DD/YYYY HH:MM:SS.
Dataset Review Date	Date after which the dataset will be reviewed by the City for utility (fit-for-purpose) and usage. Format: MM/DD/YYYY HH:MM:SS.
Comments	Specific comments related to any time-specific feature of this dataset.

Field	Definition
Start Date	The file naming convention reflects the year and month of the dataset where YYYY is the four digit year and MM is the two digit month.
End Date	
Dataset Refresh Interval	This is historical data and does not refresh.
Dataset Expiration Date	Does not expire
Dataset Review Date	None
Comments	

Dataset Definition/Format

Provide a field-by-field breakdown and definition of each record. This section acts as the formal data dictionary for an individual record.

Each abqbusesYYYYMM.txt file contains:

Field Name	Format	Description
calendardate	Numeric	The date stamp of the GPS record formatted in YYYYMMDD format.
vehiclenumber	Text	The vehicle identification number.
lat	Decimal	The latitude of the vehicle's location.
lon	Decimal	The longitude of the vehicle's location.
compassdirection	Numeric	The compass direction where 0 (zero) degrees is due north, 90 degrees is due east, 180 degrees is due south and 270 degrees is due west.
speed	Numeric	The calculated speed of the vehicle. The calculation is based on distance travelled and timespan for given distance. Speed is recorded in kilometers per hour.
msgtime	Numeric	Timestamp of the GPS record formatted in seconds after midnight.

Dataset Technical Description

Provide a technical description of the dataset. This should be a complete technical description aimed at developers and expert users who need to understand the scope, strengths and limitations of the dataset.

.txt files are provided in fixed width field format.

Columns 1-12   calendardate
Columns 14-26   vehiclenumber
Columns 28-49   lat
Columns 51-72   lon
Columns 74-89   compassdirection
Columns 91-101   speed
Columns 103-113   msgtime

Dataset Assumptions

What technical and business assumptions are implied in the creation of this dataset? Examples could include the way in which a salary figure was calculated or data that was omitted for a specific reason.

- The default polling interval of GPS data from vehicles is every 12 seconds. There are occasional deviations from the default for testing and internal purposes.

- GPS signal is degraded and false GPS trilateration calculations may be derived from buses that are parked under a canopy or parked in a garage.
- If GPS data shows that a bus moved less than 10 meters between the 12 second polling interval, it is assumed that this is GPS drift rather than physical movement. In such cases the compassdirection will not be recomputed and the last known compassdirection will be used.

Who Produced the Dataset

Which department in the City produced this dataset? Note that this might not always be in the data owner. An example of this could be a dataset that ITSD produced on behalf of EHD who owned the data.

Name	Narong Joe Saraphon
Department/Division	ABQ RIDE/IT
Email	saraphon@cabq.gov

Who Owns the Data?

Where did this data originate? Who owns the data used in this dataset? Note that this might not always be the dataset producer. An example of this could be a dataset that ITSD produced on behalf of EHD who owned the data.

The Transit Department is the data owner.

Why was the Dataset Created?

All datasets should have an explicit reason for existence and should, somehow, have value to someone. What is the perceived value that this dataset will bring?

The dataset was created to provide historical data of ABQ RIDE bus locations.

How was the Dataset Created

How was this dataset produced? Was it a manual process? An automated process? What were the main IT systems involved in producing this dataset?

The .txt files were created directly out of an SQL query from the Microsoft SQL Server Enterprise Manager.

What Similar or Related Data Should the User be Aware of?

Are there any other datasets available that may contain related or similar information? Might there be situations in which these other datasets might be a better alternative?

The historical dataset provides bus locations. In order to create an analysis of on time performance, a developer would need to also gather data from:

General Transit Feed Specification (GTFS)
http://data.cabq.gov/transit/gtfs/google_transit.zip

How Reliable are the Data?

Are there any concerns about overall data reliability? Are there any data problems that the user needs to be aware of? Are there any constraints with data accuracy? What levels of confidence with this dataset could the user reasonably assume?

Very Good.

How Well Have the Observations Been Checked?

What quality assurance steps have been performed? Sometimes, a third-party verification/audit process may also be required. If so, provide the name of the third-party who performed the verification.

Verified by the providing department.

Are there Legal Restrictions on the Access or Use of the Data?

Are there any specific legal or compliance restrictions for this data? How might this affect the way in which end users might access and use this data?

None.

Legal Disclaimer

The City's standard copyright, disclaimers and legal statements may found at http://www.cabq.gov/about/lebal. The City's data policy governing data.cabq.gov may be found at http://data.cabq.gov/policy.